These transparencies constitute a combined version of three different talks, and include a lot of supplementary material that was not actually presented in the lectures.

GENERALIZED DIFFERENTIATION THEORIES THAT SATISFY THE CHAIN RULE AND HAVE AN OPEN MAPPING PROPERTY

Héctor J. Sussmann
Rutgers University — Piscataway, NJ

and (Spring 2003)

Institut Mittag-Leffler — Djursholm, Sweden

Mathematical Control Theory Conference
Baton Rouge, Louisiana, April 10-13, 2003

(these transparencies are available at

http://www.math.rutgers.edu/~sussmann/slides.html)

$\mathcal{SET} - \mathcal{VALUED}$

$\mathcal{DIFFERENTIALS}$

Héctor J. Sussmann

Rutgers University
Piscataway, New Jersey, USA

Institut Mittag-Leffler
Djursholm, Sweden
May 20, 2003

(these transparencies are available at

http://www.math.rutgers.edu/~sussmann/slides.html)

DIFFERENTIALS OF NON–DIFFERENTIABLE MAPS

Héctor J. Sussmann
Rutgers University
Piscataway, New Jersey, USA

Brazilian Colloquium of Mathematics
IMPA
Rio de Janeiro
July 27-August 1, 2003

(these transparencies are available at

http://www.math.rutgers.edu/~sussmann/slides.html)

DO NOT WORRRY!

THE TALK

WILL BE

IN **ENGLISH**

Outline of the talk

1. Motivation: the Pontryagin Maximum Principle and generalized derivatives.
2. Categories, functors, multifunctors.
3. Statement of the extension problem.
4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. Warga derivate containers (WDCs).
6. Regular maps.
7. Generalized differential quotients (GDQs).
8. The GDQ and WDC theories are not comparable.
9. Path-integral generalized differentials (PIGDs).
10. Directional generalized differentials.

1. **Motivation: the Pontryagin Maximum Principle and generalized derivatives.**
2. Categories, functors, multifunctors.
3. Statement of the extension problem.
4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. Warga derivate containers (WDCs).
6. Regular maps.
7. Generalized differential quotients (GDQs).
8. The GDQ and WDC theories are not comparable.
9. Path-integral generalized differentials (PIGDs).
10. Directional generalized differentials.

Everybody knows that if we want nonsmooth functions such as $|x|$ to have a "derivative" at every point, then these derivatives will have to be **sets of numbers** (for functions $f : \mathbb{R} \mapsto \mathbb{R}$) or **sets of linear maps** (for functions $f : \mathbb{R}^n \mapsto \mathbb{R}^m$). For example:

> If $f : \mathbb{R} \mapsto \mathbb{R}$ is given by $f(x] = |x|$, then $f'(0)$ must be the interval $[-1, 1]$.

(Even if you don't agree with this, you should accept that $f'(0)$ contains at least 1 and -1, because the graph has two slopes at 0. Why do we actually take the convex hull? There are many reasons, and this will become clearer later. However, the only purpose of this slide is to persuade you that $f'(0)$ **must be a set**, so if you believe that $f'(0) = \{-1, 1\}$, I have won.)

OTHER EXAMPLES

If $f(x) = |x| + |x|^{3/2}\sin(1/x)$, then $f'(0)$ must be the interval $[-1,1]$.

If $f(x) = 2x + |x| + |x|^{3/2}\sin(1/x)$, then $f'(0)$ must be the interval $[1,3]$.

Homework problem: If $f(x) = x\sin(1/x)$ for $x \neq 0$, $f(0) = 0$, find $f'(0)$

Solution: Change the definition of "derivative," so that

$$f'(x) \stackrel{\text{def}}{=} \lim_{\delta \downarrow 0} \Big\{ \frac{f(x+h) - f(x)}{h} : 0 < |h| \leq \delta \Big\}.$$

Then $f'(0) = [-1,1]$.

QUESTION: what are **real mathematical definitions** that would make the above solutions correct?

REMARK: This situation is similar to that encountered by the founders of the Calculus three centuries ago (and also by our Calc I students today). They knew that the derivative of x^2 is $2x$, that the derivative of $\sin x$ is $\cos x$, that the chain rule and the Leibniz product rule hold, etc. But they did not have a good working definition of “derivative” that could be used to prove general theorems.

BUT THERE IS A BIG DIFFERENCE. Today, we know so much about mathematical rigor, that finding a rigorous mathematical definition that would make the above solutions correct is a trivial thing to do.

The problem is, we are so good at concocting rigorous definitions, that dozens of different definitions of “generalized derivative” exist. all of which agree on simple examples, but yield different results in more complicated cases.

So the really interesting question is **how do we choose the right definition?** Or, who knows, maybe we do not have to choose, and can work with all of them, as long as they satisfy some desirable properties.

WHAT WE NEED IS $\boxed{\textbf{MOTIVATION!!}}$.

We have to know what these "generalized derivatives" are supposed to be good for.

In this talk, the purpose is

To differentiate flows of nonsmooth vector fields so as to be able to prove general versions of the Pontryagin Maximum Principle.

A VERY NON-SMOOTH, HYBRID, TECHNICALLY HYBRID, MAXIMUM PRINCIPLE

Principle: A statement not yet precise enough to be a true theorem, but ready to become a theorem (in many ways) by making the terms precise.

Hybrid: Including discrete decisions, e.g., jumps.

Technically hybrid: Such that the technical conditions of different versions of the MP may be satisfied in different regions of the state space (or at different times), and our reference trajectory visits several of these regions.

Time-varying vector fields (TVVFs) and their flows

A **TVVF** on M is a (possibly partially defined) map

$$M \times \mathbb{R} \ni (x,t) \mapsto f(x,t) \in TM$$

such that $f(x,t) \in T_xM$ whenever $(x,t) \in \mathrm{Dom}(f)$.

A **trajectory** of a TVVF f on M is a locally absolutely continuous curve $\xi : I \mapsto M$, defined on a nonempty subinterval I of $\mathbb{R}$, such that

$$\boxed{(\xi(t),t) \in \mathrm{Dom}(f) \text{ and } \dot{\xi}(t) = f(\xi(t),t)}$$

for almost all $t \in I$.

Flows

If $t, s \in \mathbb{R}$, $x \in M$, we define

$$\boxed{\begin{array}{l}\Phi^f_{t,s}(x) \stackrel{\text{def}}{=} \Psi^f(x,t,s) \stackrel{\text{def}}{=} \\ \{\xi(t) : \xi \in \text{Traj}(f), \xi(s) = x\}\end{array}}\,.$$

Then Ψ^f is the **flow** of f, and the $\Phi^f_{t,s}$ are the **flow maps** of f.

Naturally, these maps are in principle set-valued, and some values may be empty, because we are not assuming existence or uniqueness of trajectories.

Differentials of flows

Suppose f is a **TVVF** on M, and $\xi_* : [a,b] \mapsto M$ is a trajectory. Let $x_* = \xi_*(a)$, $y_* = \xi_*(b)$. Then the flow map $\Phi^f_{b,a}$ sends x_* to y_*, so we may ask **what is the derivative of $\Phi^f_{b,a}$ at (x_*, y_*)?**

If $(x,t) \mapsto f(x,t)$ is of class C^1 with respect to x, measurable with respect to y, and locally bounded—together with its x-derivative—by an integrable function of t, we can compute the **classical differential** $D\Phi^f_{b,a}(x_*)$ by solving the **variational equation**

$$\dot{V}(t) = \frac{\partial f}{\partial x}\Big(\xi_*(t), t\Big) \cdot V(t)$$

with initial condition $V(a) = \text{identity}$.

Lebesgue points

Suppose f is a **TVVF** on M, and $(x,t)\in M$. We say that (x,t) is a **Lebesgue point** of f if f is defined near (x,t) and the limit of

$$\frac{1}{h}\int_{t-h}^{t+h}\sup\Big\{\|f(y,s)-f(x,t)\| : \|y-x\|\leq\delta\Big\}ds\,.$$

as $\delta\downarrow 0$ and $h\downarrow 0$ is equal to 0.

The differentiation theorems for flows in the C^1 case

Theorem I: $D\Phi^f_{b,a}(x_*)=V(b)$.

Theorem II: *If* (x_*,a) *and* (y_*,b) *are Lebesgue points of* f*, then*

$$\boxed{D\Psi^f(x_*,b,a)=f(y_*,b)dt+V(b)\Big((dx-f(x_*,a)ds\Big)}\,.$$

Differentials of Lipschitz flows

Now suppose that the TVVF $(x,t) \mapsto f(x,t)$ is Lipschitz with respect to x, measurable with respect to y, and with a Lipschitz constant locally bounded by an integrable function of t. We can compute a **Warga derivate container** $\Lambda^f_{b,a}(x_*)$ by solving the *variational differential inclusion*

$$\dot{V}(t) \in \partial_x f\Big(\xi_*(t), t\Big) \cdot V(t)$$

with initial condition $V(a) =$ identity. (Here $\partial_x f$ stands for "Clarke generalized Jacobian of f with respect to x.")

The Warga differentiation theorem for Lipschitz flows

Theorem $\mathbf{I}_{lip}$: *The set $\Lambda^f_{b,a}(x_*)$ of all $V(b)$s is a WDC of $\Phi^f_{b,a}$ at x_*.*

It would be nice of one could prove:

Theorem $\mathbf{II}_{lip}$: *If (x_*, a) and $(y_*, b))$ are Lebesgue points of f, then the set $\widehat{\Lambda}^f_{b,a}(x_*)$ of all maps $f(y_*, b)dt + V(b)\Big((dx - f(x_*, a)ds\Big)$ is a WDC of Ψ^f at (x_*, b, a).*

However:

Theorem $\mathbf{II}_{lip}$ is not true in general.

The reason is very simple: if we only assume f to be bounded by an integrable function of t, then the time-derivative of Ψ^f need not be bounded, so Ψ^f need not be Lipschitz, so Ψ^f need not have a Warga derivate container, because only Lipschitz maps have Warga derivate containers.

Question: can we make Theorem $\mathbf{II}_{lip}$ true?

We will answer this later.

A nonsmooth Max. Principle

The Maximum Principle will be a *necessary condition* for the **separation property**

$$\text{(SEP)} \qquad \boxed{\mathcal{R}^{\Sigma,\mathcal{U}}_{[a,b]}(\xi_*(a)) \cap S = \{\xi_*(b)\}}\,,$$

where $\mathcal{R}^{\Sigma}_{[a,b]}(x)$ is the Σ-reachable set from x over the interval $[a,b]$ using controls in $\mathcal{U}$, and S is some given set.

Data

(D1) a system

$$\Sigma : \quad \boxed{\dot{x} = f(x,u,t)\,,\ x \in M\,,\ u \in U\,,\ t \in \mathbb{R}}\,;$$

(D2) a "reference trajectory-control pair" (ξ_*, η_*), defined on an interval $[a,b]$;

(D3) a subset S of M s. that $y_* \stackrel{\text{def}}{=} \xi_*(b) \in S$;

(D4) a closed convex cone C in the tangent space $T_{y_*}M$;

(D5) a set $\mathcal{U}$ of functions $\eta : [a,b] \mapsto U$ (the "class of admissible controls").

(D6) a subset G (the "good set") of $[a,b]$ such that $a \in G$ and $b \in G$.

(D7) a family $\Lambda = \{\Lambda_{t,s}\}_{s,t\in[a,b],s\leq t}$ (the "derivative of the reference flow") such that each $\Lambda_{t,s}$ is a set of linear maps from $T_{\xi_*(s)}M$ to $T_{\xi_*(t)}M$.

Hypotheses

(H1) $\mathcal{U}$ contains every function $\eta : [a,b] \mapsto U$ obtained from η_* by modifying η_* on finitely many disjoint subintervals $I_1, \ldots, I_m$ of $[a,b]$ by substituting a constant control u_i for $\eta_*(t)$ on I_i.

(H2) The flow Ψ^{f_*} of the reference vector field f_* (defined by $f_*(x,t) \stackrel{\text{def}}{=} f(x,\eta_*(t),t)$) has a $\boxed{\boxed{\text{differential}}}$

$$\widehat{\Lambda}_{t,s} = \dot{\xi}_*(t)dt + \Lambda_{t,s}\Big(dx - \dot{\xi}_*(s)ds\Big)$$

at $\Big((\xi_*(s),t,s),\xi_*(t)\Big)$ whenever $s,t \in G$ and $s \leq t$.

(H3) The sets $\Lambda_{t,s}$ are compact and nonempty and satisfy the *flow identities*

$$\Lambda_{t,t} = \left\{ \mathrm{id}_{T_{\xi_*(t)}M} \right\},$$

$$\Lambda_{t,s} \circ \Lambda_{s,r} = \Lambda_{t,r} \text{ if } r, s, t \in G\,,\ r \leq s \leq t\,.$$

(H4) C is a $\boxed{\boxed{\text{tangent cone}}}$ to S at y_*.

(H5) (SEP) holds.

The $\boxed{\boxed{\text{double-boxed terms}}}$ are not yet precisely defined, and this—and *only this*—is why the principle we are stating is not yet a true theorem.

There are exactly **two** double-boxed terms:

$\boxed{\boxed{\text{differential}}}$ and $\boxed{\boxed{\text{tangent cone}}}$

Conclusion

There exist a **compatible selection**

$$L = \{L_{t,s}\}_{s,t \in G, s \le t}$$

of Λ, and covectors $\pi(t) \in T^*_{\xi_*(t)}M$, for $t \in G$, such that

(C1) *(The adjoint equation in integrated form.)* $\boxed{\pi(s) = L^\dagger_{t,s}(\pi(t))}$ whenever $s \le t$, $s, t \in G$,

NOTE: A **compatible selection** of Λ is a family $L = \{L_{t,s}\}_{s,t \in G, s \le t}$ such that $L_{t,s} \in \Lambda_{t,s}$ whenever $s \le t$ and $s, t \in G$, and $L_{t,s}L_{s,r} = L_{t,r}$ whenever $r, s, t \in G$ and $r \le s \le t$.

(C2) *(The Hamiltonian maximization condition.)*

$$\boxed{\pi(t) \cdot f(\xi_*(t), \eta_*(t), t) \geq \pi(t) \cdot f(\xi_*(t), u, t)}$$

whenever $t \in G$, $u \in U$, and t is such that $\Psi^{f(\cdot,u,\cdot)}$ is $\boxed{\boxed{\text{differentiable}}}$ at $\Big(\big(\xi_*(t), t, t\big), \xi_*(t)\Big)$ with $\boxed{\boxed{\text{differential}}}$ equal to $dx + f(\xi_*(t), u, t)(dt - ds)$.

(C3) *(The nontriviality condition.)* $\pi(b) \neq 0$.

(C4) *(The transversality condition.)* $\pi(b) \cdot c \geq 0$ whenever $c \in C$.

EXAMPLE: The *reflected brachistochrone*.

PROBLEM: Using trajectories of

$$\dot{x} = u\sqrt{|y|}\,,\;\; \dot{y} = v\sqrt{|y|}\,,\;\; u^2 + v^2 = 1\,,$$

steer $A \in \mathbb{R}^2$ to $B \in \mathbb{R}^2$ in minimum time.

The *existence* of an optimal trajectory for any given A, B is trivial (Ascoli-Arzelà).

As long as $y > 0$ or $y < 0$ the optimal trajectories are *cycloids*. (This follows from any classical version of the MP.)

With a small *ad hoc* argument, one can exclude arcs that intersect the x axis in more than point.

What is left is: arcs of the form "cycloid from A to C followed by cycloid from C to B," with C belonging to the x axis.

How does one find C?
The dynamics is not Lipschitz. So neither the classical versions of the MP nor the Lipschitz ones (Clarke, Vinter, Ioffe, Łojasiewicz Jr.) help.

Formally, it is clear that **the "adjoint vector," if it exists, will have to blow up as we cross the x axis.**

Proof: the Hamiltonian is

$$H = (p_1 u + p_2 v)\sqrt{|y|} + p_0\,.$$

Along our curve, $H \equiv 0$. So p_1 and p_2 cannot both vanish, because if they did then p_0 would be 0 as well, contradicting nontriviality. Then

$$u = \frac{p_1}{\sqrt{p_1^2 + p_2^2}} \qquad \text{and} \qquad v = \frac{p_2}{\sqrt{p_1^2 + p_2^2}}\,,$$

so the maximized Hamiltonian is

$$\mathcal{H} = \sqrt{p_1^2 + p_2^2}\sqrt{|y|} - p_0\,.$$

Since $\mathcal{H} \equiv 0$ along our trajectory, and we know that $(p_1, p_2) \neq (0,0)$, and y is not $\equiv 0$, we conclude that

$$\sqrt{p_1^2 + p_2^2} = \frac{p_0}{\sqrt{|y|}},$$

so $\sqrt{p_1^2 + p_2^2} \to +\infty$ as $y \to 0$.

So for this problem **the "adjoint vector" cannot satisfy any classical or nonsmooth "adjoint equation" whose solutions are absolutely continuous by definition.**

If there is an "adjoint equation," it will have to be allowed to have solutions that are *discontinuous* and in fact can even be **unbounded!!!!!!**.

With our version of the maximum principle there is no problem, because one can easily embed our candidate piecewise cycloid in a *flow* of piecewise cycloids, which will be smooth except when you cross the x axis. The flow will then be (classically) differentiable except at the one time τ when the reference trajectory crosses the x axis. If we form the set G by just excluding τ, we can apply our MP.

The final result is:

$$\boxed{\parbox{}{\textit{Let } R_+,\ R_- \textit{ be the radii of the rolling circles that generate the upper and lower cycloids. Then } R_1 = R_2.}}$$

1. Motivation: the Pontryagin Maximum Principle and generalized derivatives.
2. **Categories, functors, multifunctors.**
3. Statement of the extension problem.
4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. Warga derivate containers (WDCs).
6. Regular maps.
7. Generalized differential quotients (GDQs).
8. The GDQ and WDC theories are not comparable.
9. Path-integral generalized differentials (PIGDs).
10. Directional generalized differentials.

A **category** is a mathematical entity $\mathcal{C}$ consisting of the following three ingredients:

1. A class $\mathcal{O}(\mathcal{C})$ of **objects**,
2. for each pair (A, B) of objects, a set $\mathrm{Hom}_{\mathcal{C}}(A, B)$ of **morphisms** from A to B,
3. a **composition operation** $\circ_{\mathcal{C}}$, which assigns to each pair μ, ν such that $\mu \in \mathrm{Hom}_{\mathcal{C}}(A, B)$ and $\nu \in \mathrm{Hom}_{\mathcal{C}}(B, C)$ for some triple (A, B, C) of objects of $\mathcal{C}$ a morphism $\nu \circ_{\mathcal{C}} \mu$ (or just $\nu \circ \mu$) belonging to $\mathrm{Hom}_{\mathcal{C}}(A, C)$,

and such that **composition is associative** and **every object** A **admits an identity morphism** $\mathrm{id}_A \in \mathrm{Hom}_{\mathcal{C}}(A, A)$.

NOTATION: $\boxed{\mu : A \mapsto_{\mathcal{C}} B}$ (or even $\boxed{\mu : A \mapsto B}$, if it is clear what $\mathcal{C}$ is) is an alternative notation for $\mu \in \text{Hom}_{\mathcal{C}}(A, B)$.

Composition is associative means: if

$\mu : A \mapsto B$, $\nu : B \mapsto C$, and $\rho : C \mapsto D$

then $\boxed{\rho \circ (\nu \circ \mu) = (\rho \circ \nu) \circ \mu}$.

An **identity morphism** of an object A is a morphism $\iota : A \mapsto A$ such that

$$\boxed{\iota \circ \mu = \mu \text{ if } B \text{ is in } \mathcal{O}(\mathcal{C}) \text{ and } \mu : B \mapsto A,}$$

and

$$\boxed{\mu \circ \iota = \mu \text{ if } B \text{ is in } \mathcal{O}(\mathcal{C}) \text{ and } \mu : A \mapsto B,}$$

FACT: The identity morphism ι_A of A is unique.

We also require (to simplify a number of things) the following:

If $\mu : A \mapsto B$ *and* $\mu : A' \mapsto B'$
then $A = A'$ *and* $B = B'$*.*

In the definition of "category" one could do without the objects, and just take a category to consist of a class $\mathcal{M}$ of "morphisms" together with a partially defined binary operation $\circ$ on $\mathcal{M}$ called "composition."

The **objects** would then be defined as the *idempotent morphisms*, i.e., the morphisms μ such that $\mu \circ \mu = \mu$.

One could even go further and do away with the morphisms as well, and just define the category to be the composition operation.

In other words, a category would be a partially defined binary operation that satisfies certain axioms.

Since a binary operation is a collection of triples, the “morphisms” could then be defined as the entities that occur as entries of one of these triples, and the “objects” as the idempotent morphisms.

FUNCTORS

If $\mathcal{C}_1, \mathcal{C}_2$ are categories, a **functor** from $\mathcal{C}_1$ to $\mathcal{C}_2$ is a correspondence $\mathcal{F}$ that assigns to every object A of $\mathcal{C}_1$ an object $\mathcal{F}(A)$ of $\mathcal{C}_2$, and to each morphism $\mu : A \mapsto_{\mathcal{C}_1} B$ between two objects A, B of $\mathcal{C}_1$ a morphism $\mathcal{F}(\mu) : \mathcal{F}(A) \mapsto_{\mathcal{C}_2} \mathcal{F}(B)$ between the corresponding objects $\mathcal{F}(A), \mathcal{F}(B)$ of $\mathcal{C}_2$, in such a way that

a. $\boxed{\mathcal{F}(\iota_A) = \iota_{\mathcal{F}(A)}}$ whenever A is in $\mathcal{O}(\mathcal{C}_1)$,

b. $\boxed{\mathcal{F}(\nu \circ \mu) = \mathcal{F}(\nu) \circ \mathcal{F}(\mu)}$ whenever $\mu : A \mapsto_{\mathcal{C}_1} B$ and $\nu : B \mapsto_{\mathcal{C}_1} C$.

Since we are doing set-valued analysis, why not introduce set-valued functors?

A **multifunctor** from $\mathcal{C}_1$ to $\mathcal{C}_2$ is a correspondence $\mathcal{F}$ which is exactly like a functor except that if A, B are objects of $\mathcal{C}_1$ and $\mu \in \mathrm{Hom}_{\mathcal{C}_1}(A, B)$ then $\mathcal{F}(\mu)$ is a **subset** rather than a member of $\mathrm{Hom}_{\mathcal{C}_2}(\mathcal{F}(A), \mathcal{F}(B))$, and

a. $\boxed{\iota_{\mathcal{F}(A)} \in \mathcal{F}(\iota_A)}$ whenever A is in $\mathcal{O}(\mathcal{C}_1)$,

b. $\boxed{\mathcal{F}(\nu) \circ \mathcal{F}(\mu) \subseteq \mathcal{F}(\nu \circ \mu)}$ whenever A, B, C are in $\mathcal{O}(\mathcal{C}_1)$, $\mu : A \mapsto_{\mathcal{C}_1} B$, and $\nu : B \mapsto_{\mathcal{C}_1} C$.

NOTE: If we adopt the point of view that a category consists of the specification of a class of morphisms (no objects), then a multifunctor from $\mathcal{C}_1$ to $\mathcal{C}_2$ is just a correspondence $\mu \mapsto \mathcal{F}(\mu)$ that assigns to each morphism μ of $\mathcal{C}_1$ a set of morphisms $\mathcal{F}(\mu)$ of $\mathcal{C}_2$ and satisfies certain properties. The correspondence $A \mapsto \mathcal{F}(A)$ determined by $\mathcal{F}$ on the objects can then be recovered, because the objects of $\mathcal{C}_i$ are the idempotents of $\mathcal{C}_i$, and if μ is an idemptotent of $\mathcal{C}_1$, then $\mathcal{F}(\mu)$ contains one and only one idemptotent of $\mathcal{C}_2$.

A **functor** is just a multifunctor $\mathcal{F}$ such that $\text{card}(\mathcal{F}(\mu)) = 1$ for every morphism μ.

EXAMPLES OF CATEGORIES

$\boxed{\boxed{\mathcal{SVM}}}$. *(Sets and set-valued maps).* Objects: sets. Morphisms: set-valued maps. (Precisely: if A, B are sets, the members of $\mathrm{Hom}_{\mathcal{SVM}}(A,B)$ are the triples $\mu = (A,B,G)$ such that $G \subseteq A \times B$. Then A, B, G are, respectively, the *source*, *target* and *graph* of μ. **NOTATION:** $\boxed{F : A \twoheadrightarrow B}$ means $F \in \mathrm{Hom}_{\mathcal{SVM}}(A,B)$, and $\boxed{\mathrm{So}\, F,\ \mathrm{Ta}\, F,\ \mathrm{Gr}\, F,\ \mathrm{Do}\, F,\ \mathrm{Im}\, F}$, are, respectively, the source, target, graph, domain, and image of F.)

$\boxed{\boxed{\mathcal{OM}\,.}}$ *(Sets and ordinary maps.)* Objects: sets. Morphisms: everywhere defined single-valued maps. (Precisely: if A, B are sets, then the members of $\mathrm{Hom}_{\mathcal{OM}}(A,B)$ are the $\mu \in \mathrm{Hom}_{\mathcal{SVM}}(A,B)$ such that $\mathrm{Do}(\mu) = A$ and the map μ is single-valued. **NOTATION:** $\boxed{F : A \mapsto B}$ means $F \in \mathrm{Hom}_{\mathcal{OM}}(A,B)$.)

$\boxed{\boxed{\mathcal{FDRLS}}}$. *(Finite-dimensional real linear spaces and linear maps.)* Objects: finite-dimensional real linear spaces. Morphisms: linear maps. (Precisely: if A, B are FD real linear spaces, then the members of $\mathrm{Hom}_{\mathcal{FDRLS}}(A, B)$ are the $\mu \in \mathrm{Hom}_{\mathcal{OM}}(A, B)$ that are linear.)

$\boxed{\boxed{\mathcal{MC}^k\mathcal{C}^\ell}}$, if $k, \ell \in \mathbb{N} \cup \{\infty\} \cup \{\omega\}$ and $k \geq \ell$. *(Finite-dimensional manifolds of class C^k and maps of class C^ℓ.)* Objects: finite-dimensional manifolds of class C^k. The morphisms from M to M' are the maps $F : M \mapsto M'$ of class C^ℓ. **NOTATION:** $\boxed{C^\ell(M, M')}$ is the set $\mathrm{Hom}_{\mathcal{MC}^k\mathcal{C}^\ell}(M, M')$.

$\boxed{\boxed{\mathcal{PMC}^k\mathcal{C}^\ell}}$. *(Pointed finite-dimensional manifolds of class C^k and base-point preserving maps of class C^ℓ.)* Objects: pairs (M,p) such that M is a finite-dimensional manifold of class C^k and $p \in M$. The morphisms from (M,p) to (M',p') are the maps $F \in C^\ell(M,M')$ such that $F(p) = p'$. **NOTATION:** $\boxed{C^\ell((M,p),(M'p'))}$ is the set $\mathrm{Hom}_{\mathcal{PMC}^k\mathcal{C}^\ell}((M,p),(M',p'))$.

$\boxed{\boxed{\mathcal{PMC}^1\mathcal{C}^0\mathcal{D}^1}}$. *(Pointed finite-dimensional manifolds of class C^1 and base-point preserving continuous maps that are differentiable at the base point.)* Objects: same as for $\mathcal{PMC}^1$. The morphisms from (M,p) to (M',p') are the continuous maps $F : M \mapsto M'$ such that $F(p) = p'$ and F is differentiable at p. **NOTATION:** $\boxed{C^0D^1((M,p),(M'p'))}$ is the set $\mathrm{Hom}_{\mathcal{PMC}^1\mathcal{C}^0\mathcal{D}^1}((M,p),(M',p'))$.

$\boxed{\boxed{\mathcal{PMC}^1\mathcal{LIP}}}$. *(Pointed finite-dimensional manifolds of class* C^1 *and base-point preserving Lipschitz-continuous maps.)* Objects: same as for $\mathcal{PMC}^1$. The morphisms from (M,p) to (M',p') are the Lipschitz-continuous maps $F : M \mapsto M'$ such that $F(p) = p'$. **NOTATION:** $\boxed{C^{Lip}((M,p),(M'p'))}$ is the set $\mathrm{Hom}_{\mathcal{PMC}^1\mathcal{LIP}}((M,p),(M',p'))$.

$\boxed{\boxed{\mathcal{PMC}^1\mathcal{SVM}}}$. *(Pointed finite-dimensional manifolds of class* C^1 *and set-valued maps.)* Objects: same as for $\mathcal{PMC}^1$. The morphisms from (M,p) to (M',p') are the set-valued maps from M to M'.(Precisely, the 5-tuples (M,p,M',p',G), where G is a subset of $M \times M'$.)

1. Motivation: the Pontryagin Maximum
   Principle and generalized derivatives.
2. Categories, functors, multifunctors.

3. Statement of the extension problem.

4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. Regular maps.
6. Generalized differential quotients (GDQs).
7. Warga derivate containers (WDCs).
8. The GDQ and WDC theories are not comparable.
9. Path-integral generalized differentials (PIGDs).
10. Directional generalized differentials.

THE ORDINARY DIFFERENTIAL IS A FUNCTOR FROM $\mathcal{PMC}^1\mathcal{C}^1$ TO $\mathcal{FDRLS}$.

If we call this functor $\mathcal{D}_{C^1}$, then

- $\mathcal{D}_{C^1}(M,p)$ is the usual tangent space T_pM of M at p,
- if $\mu : (M,p) \mapsto_{\mathcal{PMC}^1\mathcal{C}^1} (M',p')$, then $\mathcal{D}_{C^1}(\mu)$ is the ordinary differential of μ at p, which, as is well known, is a linear map from T_pM to $T_{p'}M'$, i.e., a member of the space $\mathrm{Hom}_{\mathcal{FDRLS}}(\mathcal{D}_{C^1}(M,p), \mathcal{D}_{C^1}(M',p'))$ of all $\mathcal{FDRLS}$-morphisms from $\mathcal{D}_{C^1}(M,p)$ to $\mathcal{D}_{C^1}(M',p')$.

THE CHAIN RULE IS PRECISELY THE STATEMENT THAT $\mathcal{D}_{C^1}$ IS A FUNCTOR.

THREE VERY IMPORTANT PROPERTIES OF $\mathcal{D}_{C^1}$

THE PRODUCT RULE:

THE DIFFERENTIAL OF A PRODUCT IS THE PRODUCT OF THE DIFFERENTIALS

If, for $i = 1, 2$, $\mu_i \in C^1((M_i, p_i), (M_i', p_i'))$, then

$$\mathcal{D}_{C^1}(\mu_1 \times \mu_2) = \mathcal{D}_{C^1}(\mu_1) \times \mathcal{D}_{C^1}(\mu_2)\,.$$

LOCALITY

If $\mu, \nu \in C^1((M, p), (M', p'))$ and $\mu \equiv \nu$ on some neighborhood of p, then

$$D\mu(p) = D\nu(p)\,.$$

THE OPEN MAPPING THEOREM:

If $\mu \in C^1((M,p),(M',p'))$ and $\mathcal{D}_{C^1}(\mu)$ is surjective (i.e., $D\mu(p)$ is a surjective linear map from T_pM to $T_{p'}M'$), then μ maps neighborhoods of p in M to neighborhoods of p' in M'.

Our goal here is to discuss ways to extend $\mathcal{D}_{C^1}$ to local multifunctors defined on categories larger than $\mathcal{PMC}^1\mathcal{C}^1$ in such a way that the product rule and an open mapping theorem are still satisfied.

The **first step** is to make plenty of room for such extensions by allowing the targets of our multifunctors to take values in a category larger than $\mathcal{FDRLS}$.

NOTE: "larger" means "more morphisms." This could happen because the larger category contains more objects or, even when the class of objects is unchanged, if the are more morphisms. **Example:** $\mathcal{PMC}^1\mathcal{SVM}$ is "larger" than $\mathcal{PMC}^1\mathcal{C}^1$.

We will consider multifunctors with values in $\overline{\mathcal{FDRLS}}$, the category whose objects are the same as those of $\mathcal{FDRLS}$, but whose morphisms are **nonempty compact sets of linear maps** rather than single linear maps.

In our words, our "generalized differentials" will be nonempty compact sets of linear maps. Furthermore, **they will not be unique**, and the "chain rule" $\boxed{\boxed{\mathcal{F}(\nu) \circ \mathcal{F}(\mu) \subseteq \mathcal{F}(\nu \circ \mu)}}$ will say that

$$\boxed{\boxed{\text{The composite of two differentials is a differential of the composite map}}}$$

but will **not** say that every differential of the composite arises as a composite of differentials.

Example: If $f : \mathbb{R} \mapsto \mathbb{R}$ is given by $f(x) = 2x + |x|$, and $g(x) = \frac{1}{3}(2x - |x|)$, so that $g = f^{-1}$ (compositional inverse), then in some theories if $\Lambda \in f'(0)$ and $\Xi \in g'(0)$ then $[1, 3] \subseteq \Lambda$ and $[\frac{1}{3}, 1] \subseteq \Xi$, but $\{1\}$ is a derivative of $g \circ f$ at 0, since $(g \circ f)(x) \equiv x$.

DEFINITION: A multifunctor $\mathcal{D}$ from $\mathcal{PMC}^1\mathcal{SVM}$ to $\overline{\mathcal{FDRLS}}$ is **local** if, whenever

a. $F_i : (M,p) \mapsto_{\mathcal{PMC}^1\mathcal{SVM}} (M',p')$ for $i = 1,2$,

b. $(\mathrm{Gr}\, F_1) \cap U = (\mathrm{Gr}\, F_2) \cap U$ for some neighborhood U of (p,p') in $M \times M'$,

it follows that $\mathcal{D}(F_1) = \mathcal{D}(F_2)$.

DEFINITION: We say that a multifunctor $\mathcal{D}$ from $\mathcal{PMC}^1\mathcal{SVM}$ to $\overline{\mathcal{FDRLS}}$ **satisfies the product rule** if, whenever

$F_i : (M_i,p_i) \mapsto_{\mathcal{PMC}^1\mathcal{SVM}} (M_i',p_i')$ for $i = 1,2$,

it follows that $\mathcal{D}(F_1) \times \mathcal{D}(F_2) \subseteq \mathcal{D}(F_1 \times F_2)$.

REMARK: If you don't want your Calculus I students to get strange ideas, you can talk about the "Cartesian product rule," rather than the "product rule."

DEFINITION

A **Non-directional Generalized Differentiation Theory** (NDGDT) is a local multifunctor $\mathcal{D}$ from $\mathcal{PMC}^1\mathcal{SVM}$ to $\overline{\mathcal{FDRLS}}$ that satisfies the product rule and is an extension of $\mathcal{D}_{C^1}$.

NOTE: "$\mathcal{D}$ is an extension of $\mathcal{D}_{C^1}$" means: "if $F : (M,p) \mapsto_{\mathcal{PMC}^1\mathcal{C}^1} (M',p')$ then $\{DF(p)\} \in \mathcal{D}(F)$."

DEFINITION: If $\Lambda : X \underset{\mathcal{FDRLS}}{\longmapsto} Y$, we call Λ is **nonsingular** if every member of Λ is surjective as a linear map from X to Y

DEFINITION: If $\mathcal{D}$ is a non-directional generalized differentiation theory, we say that $\mathcal{D}$ **has the open mapping property (OMP)** if, whenever

a. $F : (M, p) \mapsto_{\mathcal{PMC}^1\mathcal{SVM}} (M', p')$

b. $\Lambda \in \mathcal{D}(F)$,

and

c. Λ is nonsingular,

it follows that F maps neighborhoods of p to neighborhoods of p' (that is, whenever $U \subseteq M$ and $p \in \operatorname{Int} U$, it follows that $p' \in \operatorname{Int} F(U)$).

1. Motivation: the Pontryagin Maximum Principle and generalized derivatives.
2. Categories, functors, multifunctors.
3. Statement of the extension problem.
4. **A trivial example of an extension: the classical differential at one point of a continuous map.**
5. Warga derivate containers (WDCs).
6. Regular maps.
7. Generalized differential quotients (GDQs).
8. The GDQ and WDC theories are not comparable.
9. Path-integral generalized differentials (PIGDs).
10. Directional generalized differentials.

One trivial extension of $\mathcal{D}_{C^1}$ that has all the desired properties is obtained by using instead of $\mathcal{PMC}^1\mathcal{C}^1$ the category $\mathcal{PMC}^1\mathcal{C}^0\mathcal{D}^1$ with the same objects as $\mathcal{PMC}^1\mathcal{C}^1$ but whose morphisms are the continuous base-point preserving maps that are differentiable at the base point.

The **open mapping theorem** says exactly the same as before (that is, if $F(p) = p'$, F is continuous near p, and $DF(p)$ exists and is surjective, then F maps neighborhoods of p to neighborhoods of p').

But the proof involves a **topological argument** in an essential way, because the result depends on (and, in fact, is elementarily equivalent to) the Brouwer fixed point theorem.

We will not pursue the study of this particular extension, because we will move on right away to a much larger extension.

THEOREM: Let Y be a real normed linear space. Write $B_Y = \{y \in Y : \|y\| \leq 1\}$. Then the following five conditions are equivalent:

a. Every $F \in C^0(B_Y, B_Y)$ has a fixed point.

b. If X is any normed space, Ω is open in X, $F \in C^0(\Omega, Y)$, $x \in \Omega$, F is Fréchet-differentiable at x, and $DF(x)$ is a surjective linear map from X to Y, then F is open at x (i.e., F maps neighborhoods of x to neighborhoods of $F(x)$).

c. If $F \in C^0(Y, Y)$, $F(0) = 0$, F is Fréchet-differentiable at 0, and $DF(0) = \mathrm{id}_Y$, then F is open at 0.

d. There does not exist a continuous retraction from B_Y to the unit sphere $\{y \in Y : \|y\| = 1\}$.

e. Y is finite-dimensional.

1. Motivation: the Pontryagin Maximum Principle and generalized derivatives.
2. Categories, functors, multifunctors.
3. Statement of the extension problem.
4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. **Warga derivate containers (WDCs).**
6. Regular maps.
7. Generalized differential quotients (GDQs).
8. The GDQ and WDC theories are not comparable.
9. Path-integral generalized differentials (PIGDs).
10. Directional generalized differentials.

WARGA DERIVATE CONTAINERS:

If $F : \mathbb{R}^m \mapsto \mathbb{R}^n$ maps 0 to 0, and Λ is a nonempty compact (not necessarily convex) subset of $\mathbb{R}^{n\times m}$, we say that Λ is a **Warga derivate container** of F at $(0,0)$, and write

$$\Lambda \in WDC(F; 0, 0)\,,$$

if for every $\delta > 0$ there exist

(a) a neighborhood U_δ of 0 in $\mathbb{R}^m$

and

(b) a sequence $\{F_j\}_{j=1}^{\infty}$ of maps of class C^1 from U_δ to $\mathbb{R}^n$

such that

(i) $DF_j(x) \in \Lambda^\delta$ whenever $x \in U_\delta$ and $j \in \mathbf{N}$

and

(ii) $F_j \to F$ uniformly on U_δ as $j \to \infty$.

Clearly, a necessary condition for F to have a WDC at $(0,0)$ is that F be Lipschitz-continuous near 0.
If this condition is satisfied, then
(a) $\partial F(0) \in WDC(F;0,0)$,
(b) $\partial F(0)$ is the smallest of all convex WDCs.

The **chain rule** and **product rule** are trivial.
One can prove an **open mapping theorem** as well.
The WDC theory extends immediately to pointed manifolds of class C^1, and is an extension of $\mathcal{D}_{C^1}$ to a local multifunctor from $\mathcal{PMC}^1\mathcal{LIP}$ to $\overline{\mathcal{FDRLS}}$ that satisfies the product rule and has the open mapping property.

WDC is a non-directional generalized differentiation theory that satisfies the OMP, contains the Clarke generalized Jacobians (and in particular the classical differentials of maps of class C^1), and assigns a nonempty set of differentials to every Lipschitz-continuous map, but does not contain the differentials at one point of continuous maps.

1. Motivation: the Pontryagin Maximum Principle and generalized derivatives.
2. Categories, functors, multifunctors.
3. Statement of the extension problem.
4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. Warga derivate containers (WDCs).
6. **Regular maps.**
7. Generalized differential quotients (GDQs).
8. The GDQ and WDC theories are not comparable.
9. Path-integral generalized differentials (PIGDs).
10. Directional generalized differentials.

REGULAR MAPS

DEFINITION. Let X, Y be metric spaces. A set-valued map $F : X \twoheadrightarrow Y$ is *regular* if for every compact subset K of X

(1) $\mathrm{Gr}\,(F\lceil K)$ is compact;

(2) there exists a sequence $\{f_j\}_{j=1}^{\infty}$ of single-valued continuous maps $f_j : K \mapsto Y$ that converges to $F\lceil K$ in the following sense:

$$\boxed{\begin{gathered}\lim_{j\to\infty} \Delta\Big(\mathrm{Gr}\,(f_j), \mathrm{Gr}\,(F\lceil K)\Big) = 0\,, \\ \text{where } \Delta(A,B) = \sup\{\,\mathrm{dist}(q,B) : q \in A\,\}.\end{gathered}}$$

(That is: "given any neighborhood U of the graph of $F\lceil K$, the graphs of the f_j are eventually inside U.") NOTE: "$\lceil$" means "restriction."

REGULARITY is a “good” property, which serves as a reasonable set-valued substitute for continuity. For example, it is invariant under composition (on either side) with continuous maps. (Actually, **a composite of regular maps is regular**.)

THE FIXED POINT THEOREM

Let K be a nonempty compact convex subset of a normed space X. Let $F : K \twoheadrightarrow K$ be a regular set-valued map. Then F has a fixed point (i.e., $(\exists x \in K)(x \in F(x))$).

THE KAKUTANI-KY FAN THEOREM

Let S be a subset of a normed space X. Let $F : S \twoheadrightarrow X$ be a set-valued map with nonempty convex values such that Gr($F\lceil K$) is compact whenever $\subseteq S$ and K is compact. Then F is regular.

In particular, *if S is nonempty compact convex and $F(S) \subseteq S$, then F has a fixed point.*

This last statement is **Ky Fan's theorem**. The finite-dimensional case is **Kakutani's theorem**. When F is single-valued we get the **Schauder fixed point theorem**, whose finite-dimensional case is the **Brouwer fixed point theorem**.

$$\begin{array}{ccc}
\text{Brouwer} & \longrightarrow & \text{Schauder} \\
\downarrow & & \downarrow \\
\text{Kakutani} & \longrightarrow & \text{KyFan}
\end{array}$$

1. Motivation: the Pontryagin Maximum Principle and generalized derivatives.
2. Categories, functors, multifunctors.
3. Statement of the extension problem.
4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. Warga derivate containers (WDCs).
6. Regular maps.
7. **Generalized differential quotients (GDQs).**
8. The GDQ and WDC theories are not comparable.
9. Path-integral generalized differentials (PIGDs).
10. Directional generalized differentials.

GENERALIZED DIFFERENTIAL QUOTIENTS

Let $m, n \in \mathbb{Z}_+$, let $F : \mathbb{R}^m \twoheadrightarrow \mathbb{R}^n$ be a set-valued map, and let Λ be a nonempty compact subset of $\mathbb{R}^{n\times m}$. We say that Λ is a **generalized differential quotient** (abbreviated "GDQ") **of** F **at** $(0,0)$, and write $\Lambda \in GDQ(F;0,0)$, if for every positive real number δ there exist U, G such that

1. U is a compact neighborhood of 0 in $\mathbb{R}^m$;
2. G is a regular set-valued map from U to the δ-neighborhood Λ^δ of Λ in $\mathbb{R}^{n\times m}$;
3. $G(x)\cdot x \subseteq F(x)$ for every $x \in U$.

The **chain rule** for GDQs is trivial:

Suppose that $F:\mathbb{R}^m \twoheadrightarrow \mathbb{R}^n$ and $F':\mathbb{R}^n \twoheadrightarrow \mathbb{R}^p$, are set-valued maps, and assume that $\Lambda \in GDQ(F;0,0)$, $\Lambda' \in GDQ(F';0,0)$.

Let $\widehat{F} = F' \circ F$, $\widehat{\Lambda} = \Lambda' \circ \Lambda$. Given δ:

- find δ_0 such that $(\Lambda')^{\delta_0} \circ \Lambda^{\delta_0} \subseteq \Lambda^{\delta}$;

- find U', $G' : U' \twoheadrightarrow (\Lambda')^{\delta_0}$ such that $G'(y).y \subseteq F'(y)$ whenever $y \in U'$;

- find U, $G : U \twoheadrightarrow \Lambda^{\delta_0}$ s. that $G(x)x \subseteq F(x)$ whenever $x \in U$;

- by shrinking U, if necessary, assume that $Lx \in U'$ whenever $x \in U$, $L \in \Lambda^{\delta_0}$.

- Let $\widehat{G}(x) = \{L'L : L \in G(x), L' \in G'(Lx)\}$.

Then if $z \in \widehat{G}(x).x$ we have $z = L'Lx$ for some $L \in G(x)$ and some $L' \in G'(Lx)$.

If $y = Lx$, then $y \in G(x)x$, so $y \in F(x)$.

Also, $y \in U'$, and $L' \in G'(y)$, so $z = L'y \in G'(y)y \subseteq F'(y)$, so $z \in F'(y)$, and then $z \in \widehat{F}(x)$.

Therefore $\widehat{G}(x).x \subseteq \widehat{F}(x)$ whenever $x \in U$.

To conclude the proof one has to show that $\widehat{G}$ is regular.

But that's easy! All we have to do is express $\widehat{G}$ as a composition of regular maps.

$$\widehat{G} = \Phi \circ \Psi \circ \Theta \circ \Gamma$$

where

$$\begin{aligned} \Gamma(x) &= G(x) \times \{x\}\,, \\ \Theta(L, x) &= (Lx, L)\,, \\ \Psi(y, L) &= G'(y) \times \{L\}\,, \\ \Phi(L', L) &= L'L\,. \end{aligned}$$

$$\boxed{\boxed{x \overset{\Gamma}{\mapsto} (L, x) \overset{\Theta}{\mapsto} (Lx, L) \overset{\Psi}{\mapsto} L', L) \overset{\Phi}{\mapsto} L'L\,.}}$$

Hence $\boxed{\boxed{\widehat{\Lambda} \in GDQ(\widehat{F}; 0, 0)}}$, as desired.

The Clarke Generalized Jacobian of a Lipschitz map is a GDQ:

PROOF: Suppose $F:\mathbb{R}^m \mapsto \mathbb{R}^n$ maps 0 to 0 and is Lipschitz-continuous near 0. Let $\Lambda = \partial F(0)$. Fix $\delta > 0$. Then Λ^δ is compact convex. Let F_j be the usual regularizations of F (obtained by convolving with an approximate identity), modified so that $F_j(0) = 0$. Then for x in some compact convex neighborhood U of 0, all the $DF_j(x)$, for $x \in U$ and j large, are in Λ^δ, and we can write

$$F_j(x) = \Big(\int_0^1 DF_j(tx)\,dt\Big)x\,.$$

So $F_j(x) \in \Lambda^\delta.x$ if j is large, and then $F(x)$ belongs to $\Lambda^\delta.x$ because $\Lambda^\delta.x$ is compact.

Given x, define

$$G^\delta(x) = \{L \in \Lambda^\delta : Lx = F(x)\}\,.$$

Then $G^\delta(x)$ is compact, convex, nonempty. Furthermore, G^δ is clearly upper semicontinuous, since the graph of G^δ is obviously compact. Therefore G^δ is regular. Since G^δ takes values in Λ^δ, and $G^\delta(x)x = F(x)$ for all x, we are done.

If $F(0) = 0$, and F is differentiable at 0 and continuous near 0 then

$$\{DF(0)\} \in GDQ(F; 0, 0).$$

This is is easy, and goes back to Botsko and Gosser.

COROLLARY (with two completely different trivial proofs) : The singleton of the ordinary differential of a map of class C^1 is a GDQ.

GDQs ARE WELL DEFINED ON MANIFOLDS:

Indeed, let us suppose that (M, p) and (M', p') are pointed manifolds of class C^1, and that $\boxed{F : (M, p) \to_{\mathcal{PMC}^1\mathcal{SVM}} (M', p')}$.

Pick charts Φ, Φ', mapping neighborhoods U, U' of p in M and of p' in M' diffeomorphically onto open balls B, B' of radius r, r' centered at 0 in Euclidean spaces of dimensions n, n', and sending p, p' to the origin. Then $\Phi' \circ F \circ \Phi^{-1}$ is a set-valued map from $\mathbb{R}^n$ to $\mathbb{R}^{n'}$.

If Λ belongs to $GDQ(\Phi' \circ F \circ \Phi^{-1}; 0, 0)$, then clearly $\boxed{D\Phi'(p')^{-1} \circ \Lambda \circ D\Phi(p)}$ is a nonempty compact set of linear maps from T_pM to $T_{p'}M'$. Furthermore, **the chain rule implies that this set does not depend on the choice of charts Φ, Φ'.**

So we can define

$$\boxed{\begin{aligned} GDQ(F; p, p') = \\ & \{ D\Phi'(p')^{-1} \circ \Lambda \circ D\Phi(p) : \\ & \Lambda \in GDQ(\Phi' \circ F \circ \Phi^{-1}; 0, 0) \} . \end{aligned}}$$

The **GDQ open mapping theorem** is trivial:

Suppose that $F : \mathbb{R}^m \twoheadrightarrow \mathbb{R}^n$ is a set-valued map, $\Lambda \in GDQ(F; 0, 0)$, and every $L \in \Lambda$ is onto.

For a surjective $L \in \mathbb{R}^{n\times m}$, let $\boxed{L^{\#} = L^{\dagger}(LL^{\dagger})^{-1}}$, so $L^{\#}$ is the *Moore-Penrose pseudoinverse* of L.

• Fix a $\delta > 0$ such that every $L \in \Lambda^{\delta}$ is onto.

• Find U, and a regular $G : U \twoheadrightarrow \Lambda^{\delta}$, such that $G(x)x \subseteq F(x)$ whenever $x \in U$. Assume U is the closed r-ball centered at 0.

• For $x \in U$, let $G^{\#}(x) = \{L^{\#} : L \in G(x)\}$. Then $G^{\#}$ is a regular map from U to $\mathbb{R}^{m\times n}$.

• Let C be such that $\|M\| \leq C$ whenever $x \in U$ and $M \in G^{\#}(x)$.

• Given y such that $C\|y\| \leq r$, let μ_y be the set-valued map $U \ni x \mapsto G^{\#}(x)y$. Then μ_y maps U to U and is regular. So μ_y has a fixed point x. Then $x \in G^{\#}(x)y$, so $x = My$ for some $M \in G^{\#}(x)$. Then $M = L^{\#}$ for some $L \in G(x)$. So $x = L^{\#}y$, and then $Lx = LL^{\#}y = y$, so $y \in G(x)x \subseteq F(x)$. **END**

Locality is trivially true.

The **GDQ product rule** is trivial as well.

Hence

GDQ is a non-directional generalized differentiation theory that satisfies the open mapping property and contains the classical differentials at one point of continuous maps as well as the Clarke generalized Jacobians.

GDQs of Lipschitz flows

Suppose that $M \times \mathbb{R} \ni (x,t) \mapsto f(x,t)$ is Lipschitz with respect to x, measurable with respect to y, and with a Lipschitz constant locally bounded by an integrable function of t. Let $\xi_* : [a,b] \mapsto M$ be a trajectory of f. Solve the *variational differential inclusion*

$$\dot{V}(t) \in \partial_x f\Big(\xi_*(t), t\Big) \cdot V(t)$$

with initial condition $V(a) =$ identity.

Theorem $\mathbf{I}_{lip}$: *The set $\Lambda^f_{b,a}(x_*)$ of all $V(b)$s is a GDQ of $\Phi^f_{b,a}$ at (x_*, y_*).*

Theorem $\mathbf{II}_{lip}$: *If (x_*, a) and $(y_*, b))$ are Lebesgue points of f, then the set $\widehat{\Lambda}^f_{b,a}(x_*)$ of all maps $f(y_*, b)dt + V(b)\Big((dx - f(x_*, a)ds\Big)$ is a GDQ of Ψ^f at $((x_*, b, a), y_*)$.*

1. Motivation: the Pontryagin Maximum Principle and generalized derivatives.
2. Categories, functors, multifunctors.
3. Statement of the extension problem.
4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. Warga derivate containers (WDCs).
6. Regular maps.
7. Generalized differential quotients (GDQs).
8. **The GDQ and WDC theories are not comparable.**
9. Path-integral generalized differentials (PIGDs).
10. Directional generalized differentials.

An example of a GDQ that is not a WDC

Let $f : \mathbb{R} \mapsto \mathbb{R}$ be given by

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{if} \quad x \neq 0 \\ 0 & \text{if} \quad x = 0 \end{cases} .$$

Then $f(x) = G(x)x$, where

$$G(x) = \begin{cases} \left\{\sin \frac{1}{x}\right\} & \text{if} \quad x \neq 0 \\ [-1, 1] & \text{if} \quad x = 0 \end{cases} .$$

Furthermore, G is regular, because it is usc and has compact convex nonempty values. So

$$[-1, 1] \in GDQ(f, 0, 0) .$$

An example of a WDC which is not a GDQ

Let $f:\mathbb{R}\mapsto\mathbb{C}$ be given by

$$f(x)=\int_0^x e^{i/t}dt\,.$$

Then f is Lipschitz, and f is in fact of class C^1 on $\mathbb{R}\backslash\{0\}$.

The circle $\mathbb{S}^1=\{z\in\mathbb{C}:\|z\|=1\}$ is a WDC of f at $(0,0)$, but is not a GDQ. (Reason: $f(x)=O(x^2)$ as $x\to 0$. Therefore if $G(x)x=f(x)$ then $G(x)=\frac{f(x)}{x}$ for $x\neq 0$, and then $G(x)=O(|x|)$ as $x\to 0$.)

We have seen examples of (i) a WDC which is not a GDQ and (ii) a GDQ which is not a WDC. Therefore

$$\boxed{\boxed{\textbf{The theories WDC and GDQ are not comparable.}}}$$

Several extensions of WDC have been proposed (e.g., Halkin's "screens," as well as HJS's "semidifferentials" and "multidifferentials") which are not comparable with GDQ either. (Halkin also proposed a theory of "fans" which goes in the GDQ direction but is slightly more restrictive.)
Here we propose a theory that contains all the above. We call the differentials of this theory **path-integral generalized differentials**

1. Motivation: the Pontryagin Maximum Principle and generalized derivatives.
2. Categories, functors, multifunctors.
3. Statement of the extension problem.
4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. Warga derivate containers (WDCs).
6. Regular maps.
7. Generalized differential quotients (GDQs).
8. The GDQ and WDC theories are not comparable.
9. **Path-integral generalized differentials (PIGDs).**
10. Directional generalized differentials.

BASIC IDEA: Let us write $\mathbf{B}_n(r)$ to denote the closed ball $\mathbf{B}_n(r) \stackrel{\text{def}}{=} \{\, x \in \mathbb{R}^n : \|x\| \le r \,\}$.

Use $\mathcal{A}_n(r)$ to denote the set of all Lipschitz paths $\alpha : [0,1] \mapsto \mathbb{R}^n$ such that $\alpha(0) = 0$ and $\|\dot{\alpha}(t)\| \le r$ for almost all $t \in [0,1]$.

We want to think of $\mathcal{A}_n(r)$ as a "blowup" of $\mathbf{B}_n(r)$. The map $\mathcal{P}_{n,r} : \mathcal{A}_n(r) \mapsto \mathbf{B}_n(r)$ such that $\mathcal{P}_{n,r}(\alpha) = \alpha(1)$ for $\alpha \in \mathcal{A}_n(r)$ is onto, and we can think of maps from $\mathbf{B}_n(r)$ to some space as constructed starting from maps defined on $\mathcal{A}_n(r)$.

Actually,

$\boxed{\textit{the inverse } \mathcal{P}_{n,r}^{-1} : \mathbf{B}_n(r) \twoheadrightarrow \mathcal{A}_n(r) \textit{ is a regular set-valued map.}}$

PROOF: the graph of $\mathcal{P}_{n,r}^{-1}$ is the set

$$\{(x, \alpha) : x \in \mathbf{B}_n(r) \wedge \alpha \in \mathcal{A}_n(r) \wedge \alpha(1) = x\},$$

which is clearly compact. Furthermore, each set $\mathcal{P}_{n,r}^{-1}(x)$ is convex and nonempty.

We will consider set-valued maps $F:\mathbb{R}^n \mapsto \mathbb{R}^m$ that contain maps $f_r:\mathbf{B}_n(r)\mapsto\mathbb{R}^m$ of the form $\Phi_r\circ\mathcal{P}_{n,r}^{-1}$, where $\Phi_r:\mathcal{A}_n(r)\mapsto\mathbb{R}^m$ is of the form

$$\Phi_r(\alpha)=\int_0^1 h_\alpha(t)\dot\alpha(t)dt+v_\alpha\,, \qquad (1)$$

and v_α is a "small error."

The map $\alpha\mapsto(h_\alpha,v_\alpha)$ will take values in the product $C^0([0,1],\mathbb{R}^{m\times n})\times\mathbb{R}^n$, and will in fact be allowed to be set-valued, but will be required to be regular.

This suffices to imply that **the set-valued map Φ_r is regular too**.

PROOF: it suffices to see that the map $(h,\alpha) \mapsto \int_0^1 h(t)\dot\alpha(t)dt$ is continuous on $\Sigma_r \stackrel{\text{def}}{=} C^0([0,1],\mathbb{R}^{m\times n}) \times \mathcal{A}_n(r)$. If (h_j,α_j) converge to (h,α) in Σ_r, then

$$\int_0^1 h_j\dot\alpha_j - \int_0^1 h\dot\alpha = \int_0^1 (h_j-h)\dot\alpha_j + \int_0^1 h(\dot\alpha_j-\dot\alpha)\,.$$

Clearly, $\int_0^1 (h_j-h)\dot\alpha_j \to 0$. Given $\varepsilon > 0$, find a $g \in C^1$ such that $\|h-g\|_{sup} \le \varepsilon$. Then, if we let $\beta_j = \alpha_j - \alpha$, we have

$$\int_0^1 g(\dot\alpha_j-\dot\alpha) = g(1)\beta_j(1) - g(0)\beta_j(0) - \int_0^1 \dot g(t)\beta_j dt,$$

which goes to zero, because $\beta_j \to 0$ uniformly. And, finally, $\|\int_0^1 (h-g)(\dot\alpha_j-\dot\alpha)\| \le 2\varepsilon r$.

Finally, if we want Λ to be a "differential" of F, we will require that the h_α's that occur in (1) be "Λ-valued modulo a small error" (i.e., Λ^δ-valued with arbitrarily small δ), and that the error v_α be $o(r)$, i.e., bounded by δr with arbitrarily small δ.

SUMMARIZING: We will look at set-valued maps $F:\mathbb{R}^n \twoheadrightarrow \mathbb{R}^m$ whose graphs contain graphs of maps $f_r:\mathbf{B}_n(r)\twoheadrightarrow \mathbb{R}^m$ of the form $\Phi_r \circ \mathcal{P}_{n,r}^{-1}$, where $\Phi_r:\mathcal{A}_n(r)\twoheadrightarrow \mathbb{R}^m$ is of the form

$$\Phi_r(\alpha)=\Big\{v+\int_0^1 h(t)\dot{\alpha}(t)dt:(h,v)\in G_r(\alpha)\Big\}, \quad (2)$$

where G_r is a regular set-valued map from $\mathcal{A}_n(r)$ to $C^0([0,1],\Lambda^\delta)\times \mathbf{B}_m(\delta r)$.

PIGDs

DEFINITION: Let $n, m \in \mathbb{Z}_+$, $F : \mathbb{R}^n \twoheadrightarrow \mathbb{R}^m$, and let Λ be a nonempty compact subset of $\mathbb{R}^{m\times n}$. We say that Λ is a **path-integral generalized differential** of F at $(0,0)$, and write $\boxed{\Lambda \in PIGD(F;0,0)}$, if for every $\delta > 0$ there exists $R \in]\,0,\infty\,[$ such that for every $r \in]\,0,R]$ there exists a regular set-valued map

$$G : \mathcal{A}_n(r) \twoheadrightarrow C^0([0,1],\mathbb{R}^{m\times n}) \times \mathbb{R}^m$$

such that

(1) $h(t) \in \Lambda^\delta$ and $\|v\| \leq \delta r$ whenever $\alpha \in \mathcal{A}_n(r)$, $(h,v) \in G(\alpha)$, $t \in [0,1]$,

(2) $v + \int_0^1 h(t)\dot{\alpha}(t)dt \in F(\alpha(1))$ whenever $\alpha \in \mathcal{A}_n(r)$, $(h,v) \in G(\alpha)$.

The **chain rule for PIGDs** is trivial: Suppose that $F:\mathbb{R}^m \twoheadrightarrow \mathbb{R}^n$, $F':\mathbb{R}^n \twoheadrightarrow \mathbb{R}^p$, are set-valued maps, and that $\Lambda \in PIGD(F;0,0)$ and $\Lambda' \in PIGD(F';0,0)$. Let $\widehat{F} = F' \circ F$, $\widehat{\Lambda} = \Lambda' \circ \Lambda$. Given δ:

- define

$$\begin{aligned} C &= \sup\{\|L\| : L \in \Lambda^{\delta}\}\,, \\ C' &= \sup\{\|L\| : L \in (\Lambda')^{\delta}\}\,. \end{aligned}$$

- find δ_0 such that

$$\begin{aligned} 0 < \delta_0 &\leq \delta\,, \\ \delta_0(C+\delta_0) + C'\delta_0 &\leq \delta\,, \\ (\Lambda')^{\delta_0} \circ \Lambda^{\delta_0} &\subseteq \Lambda^{\delta}\,; \end{aligned}$$

- find R' and for each $r' \in]0,R']$ a regular map $G'_{r'} : \mathcal{A}_n(r') \twoheadrightarrow C^0([0,1],(\Lambda')^{\delta_0} \times \mathbb{R}^p$ such that

  $$w + \int_0^1 k\dot{\beta} \in F'(y) \quad \text{and} \quad \|w\| \leq \delta_0 r'$$

  whenever β belongs to $\mathcal{A}_n(r)$, $y = \beta(1)$, and $(k,w) \in G'_{r'}(\beta)$,

- find R and for each $r \in]0, R]$ a regular map $G_r : \mathcal{A}_m(r) \twoheadrightarrow C^0([0,1], (\Lambda)^{\delta_0} \times \mathbb{R}^n$ such that

$$v + \int_0^1 h(t)\dot{\alpha}(t)dt \in F(x) \quad \text{and} \quad \|v\| \leq \delta_0 r$$

whenever $\alpha \in \mathcal{A}_m(r)$, $x = \alpha(1)$, $(h, v) \in G(\alpha)$,

- for $r \in]0, R]$, $\alpha \in \mathcal{A}_m(r)$, $(h, v) \in G_r(\alpha)$, define

$$\beta_{r,\alpha,h,v}(t) = \int_0^t \Big(v + h(s)\dot{\alpha}(s)\Big)ds\,.$$

Then $\beta_{r,\alpha,h,v} \in \mathcal{A}_m(r')$ if $r' = (C + \delta_0)r$.

- Let $\widehat{G}_r(\alpha)$ be the set of all pairs

$$\Big(kh, w + \Big(\int_0^1 k(t)dt\Big)v\Big)\,,$$

for all pairs (h, v) belonging to $G_r(\alpha)$ and all $(k, w) \in G'_{(C+\delta_0)r}(\beta_{r,\alpha,h,v})$. (This makes sense if $(C + \delta_0)r \leq R'$.)

Then if

$$z = u + \int_0^1 \ell(t)\dot\alpha(t)dt$$

for some $(\ell, u) \in \widehat{G}(\alpha)$, we can write

$$u = w + (\int_0^1 k(t)dt)v\,,\ \ \ell(t) = k(t)h(t)\,,$$

and get

$$\begin{aligned} z &= w + (\int_0^1 k(t)dt)v + \int_0^1 k(t)h(t)\dot\alpha(t)dt \\ &= w + \int_0^1 (k(t)(v + h(t)\dot\alpha(t))dt \\ &= w + \int_0^1 k(t)\dot\beta_{r,\alpha,h,v}(t)dt\,. \end{aligned}$$

If $y = \beta_{\alpha,h,v}(1)$, then $y \in F(x)$. Then $z \in F'(y)$. So $z \in \widehat{F}(x)$. Furthermore,

$$\|w\| \le \delta_0(C + \delta_0)r + C'\delta_0 r \le \delta r\,.$$

Hence the maps $\widehat{G}_r$, constructed for $r \in]0, \widehat{R}]$, where

$$\widehat{R} = \min\Big(R, \frac{R'}{C+\delta_0}\Big)\,,$$

satisfy all the desired properties, provided that every $\widehat{G}_r$ is regular.
To conclude the proof one has to show that the maps $\widehat{G}_r$ are regular. But that's easy! All we have to do is express $\widehat{G}$ as a composition of regular maps, as we did for the GDQ case. Hence $\boxed{\boxed{\widehat{\Lambda} \in PIGD(\widehat{F}; 0, 0)}}$, as desired.

PIGDs ARE WELL DEFINED ON MANIFOLDS:

Indeed, let us suppose that (M, p) and (M', p') are pointed manifolds of class C^1, and that $\boxed{F : (M, p) \rightarrow_{\mathcal{PMC}^1\mathcal{SVM}} (M', p')}$. Pick charts Φ, Φ', mapping neighborhoods U, U' of p in M and of p' in M' diffeomorphically onto open balls B, B' of radius r, r' centered at 0 in Euclidean spaces of dimensions n, n', and sending p, p' to the origin. Then $\Phi' \circ F \circ \Phi^{-1}$ is a set-valued map from $\mathbb{R}^n$ to $\mathbb{R}^{n'}$.

If Λ belongs to $PIGD(\Phi' \circ F \circ \Phi^{-1}; 0, 0)$, then clearly $\boxed{D\Phi'(p')^{-1} \circ \Lambda \circ D\Phi(p)}$ is a nonempty compact set of linear maps from T_pM to $T_{p'}M'$. Furthermore, **the chain rule implies that this set does not depend on the choice of charts Φ, Φ'.**

So we can define

$$\boxed{\begin{aligned} PIGD(F;p,p') \quad = \quad & \{\, D\Phi'(p')^{-1} \circ \Lambda \circ D\Phi(p) : \\ & \Lambda \in GDQ(\Phi' \circ F \circ \Phi^{-1}; 0,0)\,\}\,. \end{aligned}}$$

The **open mapping theorem for PIGDs** is trivial:

Suppose that $F : \mathbb{R}^m \twoheadrightarrow \mathbb{R}^n$ is a set-valued map, $\Lambda \in PIGD(F; 0, 0)$, and every $L \in \Lambda$ is onto.

For an onto $L \in \mathbb{R}^{m\times n}$, let $\boxed{L^{\#} = L^{\dagger}(LL^{\dagger})^{-1}}$, so $L^{\#}$ is the *Moore-Penrose pseudoinverse* of L.

- Fix a $\bar{\delta} > 0$ such that every $L \in \Lambda^{\bar{\delta}}$ is onto.
- Let $C = \max\{\|L^{\#}\| : L \in \Lambda^{\bar{\delta}}\}$.
- Pick δ such that $0 < \delta < \min\left(\frac{1}{C}, \bar{\delta}\right)$.
- Find r, and a regular

$$G : \mathcal{A}_n(r) \twoheadrightarrow C^0([0,1], \mathbb{R}^{m\times n}) \times \mathbb{R}^m$$

  such that $v + \int_0^1 h(t)\dot{\alpha}(t)dt \in F(\alpha(1))$ and $\|v\| \le \delta r$ whenever $\alpha \in \mathcal{A}_n(r)$, $(h, v) \in G(\alpha)$, and $h(t) \in \Lambda^{\delta}$ whenever $t \in [0, 1]$, $\alpha \in \mathcal{A}_n(r)$, and $(h, v) \in G(\alpha)$.

- Given $y \in \mathbb{R}^m$ such that $C\|y\| \leq (1 - C\delta)r$, we want to find an $x \in \mathbf{B}_n(r)$ such that $y \in F(x)$. For this purpose, it suffices to find $\alpha \in \mathcal{A}_n(r)$, $(h, v) \in G(\alpha)$ such that

$$v + \int_0^1 h(t)\dot{\alpha}(t)dt = y\,.$$

So we want to solve $\boxed{\int_0^1 h(t)\dot{\alpha}(t)dt = y - v}$, for which it suffices to find α, h, v with the property that $\boxed{\beta_{\alpha,h,v,y} = \alpha}$ where

$$\beta_{\alpha,h,v,y}(t) = \int_0^t h(s)^{\#}(y - v)ds\,. \qquad (3)$$

(Indeed, (3) implies

$$\dot{\beta}_{\alpha,h,v,y}(t) = h(t)^{\#}(y - v) \qquad \text{a.e.},$$

from which it follows that $h(t)\dot{\beta}_{\alpha,h,v,y}(t) = y - v$ (because $h(t)h(t)^{\#} = \mathrm{id}_{\mathbb{R}^n}$), and then we get $h(t)\dot{\alpha}(t) = y - v$, so $\int_0^1 h(t)\dot{\alpha}(t)\,dt = y - v$, as desired.)

Let μ_y be the set-valued map that sends a path $\alpha \in \mathcal{A}_n(r)$ to the set

$$\left\{ \beta_{\alpha,h,v,y} : (h,v) \in G(\alpha) \right\} \subseteq C^0([0,1],\mathbb{R}^n)\,.$$

Then μ_y maps $\mathcal{A}_n(r)$ to $\mathcal{A}_n(r)$ and is regular. So μ_y has a fixed point α, and we are done.

(NOTE: Part of this argument is based on an old idea that Waszewski used in his work on the domain of an impicitly defined function.)

Warga Derivate Containers of Lipschitz maps are PIGDs:

Indeed, suppose $F : \mathbb{R}^m \mapsto \mathbb{R}^n$ maps 0 to 0 and the map is Lipschitz-continuous near 0. Let $\Lambda \in WDC(F; 0, 0)$. Fix $\delta > 0$. Then Λ^δ is compact.

Let F_j be maps of class C^1 that are defined on a ball $\mathbf{B}_m(R) = \{x \in \mathbb{R}^m : \|x\| \leq R\}$ and are such that $F_j(0) = 0$, $DF_j(x) \in \Lambda^\delta$ for all $x \in \mathbf{B}_m(R)$ and all j, and $F_j \to F$ uniformly on $\mathbf{B}_m(R)$.

Given a number $r \in]0, R]$, pick a $j \in \mathbf{N}$ such that $\|F_j(x) - F(x)\| \leq \delta r$ for all $x \in \mathbf{B}_m(r)$, and define

$$\begin{aligned} h_\alpha(t) &= DF_j(\alpha(t)), \\ v_\alpha &= F(\alpha(1)) - F_j(\alpha(1)), \end{aligned}$$

for $\alpha \in \mathcal{A}_n(r)$.

Then

$$\begin{aligned} F(\alpha(1)) &= v_\alpha + F_j(\alpha(1)) \\ &= v_\alpha + \int_0^1 \frac{d}{dt}\Big(F_j(\alpha(t))\Big)dt \\ &= v_\alpha + \int_0^1 DF_j(\alpha(t))\dot{\alpha}(t)dt \\ &= v_\alpha + \int_0^1 h_\alpha(t)\dot{\alpha}(t)dt\,. \end{aligned}$$

END

GDQs are PIGDs:

Proof. Suppose that $F : \mathbb{R}^m \twoheadrightarrow \mathbb{R}^n$, and Λ belongs to $GDQ(F; 0, 0)$. Fix $\delta > 0$. Let $G : \mathbf{B}_n(r) \twoheadrightarrow \Lambda^\delta$ be a regular set-valued map such that the inclusion $G(x)x \subseteq F(x)$ is true whenever $x \in \mathbf{B}_n(r)$. Define

$$\mathcal{G}(\alpha) = \{(\hat{L}, 0) : L \in G(\alpha(1))\} ,$$

where, if $L \in \mathbb{R}^{n\times m}$, $\hat{L}$ is the constant matrix-valued function $[0, 1] \ni t \mapsto L \in \mathbb{R}^{n\times m}$.
Then, if $(h, v) \in \mathcal{G}(\alpha)$, and we let $x = \alpha(1)$, we have $h(t) \equiv L$ for some $L \in G(x)$, and then $v + \int_0^1 h(t)\dot{\alpha}(t)dt = Lx$, so

$$v + \int_0^1 h(t)\dot{\alpha}(t)dt \in G(x)x \subseteq F(x) .$$

Furthermore, $\mathcal{G}(\alpha)$ is regular, because

$$\mathcal{G}(\alpha) = \boldsymbol{\Theta} \circ G \circ \mathcal{P}_{m,r}$$

where $\boldsymbol{\Theta}$ is the map

$$\mathbb{R}^{n\times m} \ni L \mapsto (\hat{L}, 0) \in C^0([0,1], \mathbb{R}^{n\times m}) \times \mathbb{R}^n \,.$$

END

The **PIGD product rule** and the **locality property** are trivial as well. Hence

PIGD is a non-directional generalized differentiation theory that satisfies the open mapping property and contains GDQ and WDC. In addition, PIGD contains all the other theories (e.g., semidifferentials and multidifferentials) mentioned but not discussed in detail in this talk.

This concludes the EASY part of the talk, in which all the categories that were sources of our extensions had as objects those of $\mathcal{PMC}^1\mathcal{C}^1$. (That is, we extended $\mathcal{D}_{C^1}$ to larger classes of maps, but without enlarging the class of objects.)

1. Motivation: the Pontryagin Maximum Principle and generalized derivatives.
2. Categories, functors, multifunctors.
3. Statement of the extension problem.
4. A trivial example of an extension: the classical differential at one point of a continuous map.
5. Warga derivate containers (WDCs).
6. Regular maps.
7. Generalized differential quotients (GDQs).
8. The GDQ and WDC theories are not comparable..
9. Path-integral generalized differentials (PIGDs).
10. **Directional generalized differentials.**

It would be desirable to do similar things for **directional derivatives**, i.e., derivatives along subsets.

That is, one should consider categories whose objects are triples (M, S, p) consisting of a manifold M of class C^1, a subset S of M such that $p \in S$, and a point p of M, and such that the morphisms from (M, S, p) to (M', S', p') are set-valued maps from S to S'.

When we try to do that, all kinds of technical issues arise that need to be addressed.

Some are trivial, some seem truly hard.

Among the trivial issues are:

1. Shall we take **all** subsets S or only some "nice" sets?
2. What are the "tangent objects" (i.e., the "object" parts of our functors) of the triples (M, S, p). Should they still be the spaces T_pM, or should they be "tangent cones" in some sense to the sets S?
3. And if we choose the latter possibility, what kind of tangent cone?

4. If we are going to work with set-valued maps from S to S', and try to define their "differentials" at (p, p'), then it is clear why we need S, S', p and p'. But why do we need M and M'? Shouldn't we just formulate the theory in terms of "pointed sets," i.e. pairs (S, p), without bringing in the "ambient manifolds" M at all? But if we do this we would still need some extra structure on S that enables us to talk about "derivatives at p" of at least *some* maps defined on S. How shall such a structure be defined intrinsically, without embedding S in a manifold M?

These questions are trivial, not in the sense that they have an obvious unique answer, but in the sense that they have lots of different equally reasonable answers, and one can list them all and feel free to choose one.

But all these answers seem to come with a price. For example, if one insists on having too many sets S, then proving the chain rule becomes harder.

LOCAL DIFFCONES

DEFINITION: A **local diffcone of class C^k** is a triple (M, S, p) such that

a. M is a manifold of class C^k,

b. S is a subset of M,

c. $p \in S$,

d. there exist Φ, U, r, C such that

  d1. U is a neighborhood of p in M,

  d2. $r > 0$,

  d3. Φ is a diffeomorphism of class C^k from U to the closed ball $\mathbf{B}_n(r)$,

  d4. $\Phi(p) = 0$,

  d5. C is a closed convex cone in $\mathbb{R}^n$,

  d6. $\Phi(S \cap U) = \mathbf{B}_n(r) \cap C$.

TRIVIAL REMARK: If (M, S, p) is a local diffcone of class C^1, then all existing notions of "tangent cone of S at p" (e.g., Bouligand, Clarke) agree, and yield a closed convex cone in T_pM that certainly deserves to be called **the** tangent cone of (M, S, p).

In particular: if we want to define "directional generalized differentiation theories" to be multifunctors on the category $\mathcal{LDCC}^1$ of local diffcones of class C^1, then it is clear how to define the "object-to-object" part of the multifunctor: to each object (M, S, p) of $\mathcal{LDCC}^1$ we assign the tangent cone to Sm at p.

DISGUSTINGLY PEDANTIC AND TOTALLY USELESS REMARK, FOR BOURBAKI FANS ONLY *(all others please skip this)*: One can define the notion of "diffcone" intrinsically, without referring to the ambient manifold M at all. (Idea: a *diffcone of class* C^k is a pointed topological space (S,p) endowed with a sheaf $\mathcal{S}$ of rings of real-valued functions such there exist

(a) a neighborhood U of p in S,

(b) a neighborhood V of 0 in $\mathbb{R}^n$ for some n, and

(c) a closed convex cone C in $\mathbb{R}^n$,

such that the ringed spaces $(U, \mathcal{S}\lceil U)$ and $(C, C^k(C)\lceil V)$ are isomorphic by means of a map $\Phi : U \mapsto V$ such that $\Phi(p) = 0$, where $C^k(C)$ is the sheaf of restrictions to C of real-valued functions of class C^k defined on open subsets of $\mathbb{R}^n$.)

PATH-INTEGRAL GENERALIZED DIFFERENTIALS ON LOCAL DIFFCONES

BASIC IDEA:

Let C be a closed convex cone in $\mathbb{R}^n$. Recall that $\mathbf{B}_n(r) \stackrel{\text{def}}{=} \{ x \in \mathbb{R}^n : \|x\| \leq r \}$. Set

$$C(r) \stackrel{\text{def}}{=} C \cap \mathbf{B}_n(r)\,.$$

Use $\mathcal{A}(C(r))$ to denote the set of all Lipschitz paths $\alpha : [0,1] \mapsto \mathbb{R}^n$ such that $\alpha(0) = 0$ and $\dot{\alpha}(t) \in C(r)$ (i.e., $\dot{\alpha}(t) \in C$ and $\|\dot{\alpha}(t)\| \leq r$) for almost all $t \in [0,1]$.

We want to think of $\mathcal{A}(C(r))$ as a "blowup" of $C(r)$.

The map $\mathcal{P}_{C,r}$ such that $\boxed{\mathcal{P}_{C,r}(\alpha) = \alpha(1)}$ for $\alpha \in \mathcal{A}(C(r))$ is onto, and we can think of maps from $C(r)$ to some space as constructed starting from maps defined on $\mathcal{A}(C(r))$.

Actually, the inverse

$$\mathcal{P}_{C,r}^{-1} : C(r) \twoheadrightarrow \mathcal{A}(C(r))$$

is a regular set-valued map. (Proof: the graph of $\mathcal{P}_{C,r}^{-1}$ is the set

$$\{(x,\alpha) : x \in C(r) \wedge \alpha \in \mathcal{A}(C(r)) \wedge \alpha(1) = x\},$$

which is clearly compact. Furthermore, each set $\mathcal{P}_{C,r}^{-1}(x)$ is convex and nonempty.)

We will consider set-valued maps F that contain maps $f_r : C(r) \mapsto \mathbb{R}^m$ of the form $\Phi_r \circ \mathcal{P}_{C,r}^{-1}$, where $\Phi_r : \mathcal{A}(C(r)) \mapsto \mathbb{R}^m$ is of the form

$$\Phi_r(\alpha) = \int_0^1 h_\alpha(t)\dot{\alpha}(t)dt + v_\alpha\,, \qquad (4)$$

and v_α is a "small error."

The map $\alpha \mapsto (h_\alpha, v_\alpha)$ will take values in the product $C^0([0,1], \mathbb{R}^{m\times n}) \times \mathbb{R}^n$, and will in fact be allowed to be set-valued, but will be required to be regular.

This suffices to imply that the set-valued map Φ_r is regular too.

Finally, if we want Λ to be a "differential" of F, we will require that the h_α's that occur in (1) be "Λ-valued modulo a small error" (i.e., Λ^δ-valued with arbitrarily small δ), and that the error v_α be $o(r)$, i.e., bounded by δr with arbitrarily small δ.

Let $n, m \in \mathbb{Z}_+$, let C, C' be closed convex cones in $\mathbb{R}^n$, $\mathbb{R}^m$, respectively, let $F : C \twoheadrightarrow C'$, and let Λ be a nonempty compact subset of $\mathrm{Lin}(C, C')$. We say that Λ is a **path-integral generalized differential** of F at $(0,0)$, and write $\Lambda \in PIGD(F; 0, 0)$, if for every $\delta > 0$ there exists $R \in]\,0, \infty\,[$ such that for every $r \in]\,0, R]$ there exists a regular set-valued map

$$G : \mathcal{A}(C(r)) \twoheadrightarrow C^0([0,1], \mathbb{R}^{m \times n}) \times \mathbb{R}^m$$

such that

(1) $h(t) \in \Lambda^\delta$ and $\|v\| \leq \delta r$ whenever $\alpha \in \mathcal{A}(C(r))$, $(h, v) \in G(\alpha)$, $t \in [0,1]$,

(2) $v + \int_0^1 h(t)\dot{\alpha}(t)dt \in F(\alpha(1))$ whenever $\alpha \in \mathcal{A}_n(r)$, $(h, v) \in G(\alpha)$.

If one could prove the chain rule, then one could define PIGDs on local diffcones of class C^1.

It then turns out that the locality property and the product rule are trivial.

Furthermore, the **open mapping theorem**, properly reformulated, can also be proved.

THEOREM O: *Let* (M, S, p), (M', S', p'), *be diffcones of class* C^1, *let* $F : S \twoheadrightarrow S'$, *and let* $\Lambda \in PIGD(F, p, p')$. *Let* C, C' *be the tangent cones of* (M, S, p), (M', S', p'), *respectively. Assume that* v *is an interior point of* C' *(as a closed convex cone) and* D *is a closed convex cone such that* $D \subseteq C'$ *and* $v \in \mathrm{Int}_{C'} D$. *Assume that* $D \subseteq LC$ *for every* $L \in \Lambda$. *Let* $\xi : [0, \bar{\varepsilon}] \mapsto S'$ *be, for some positive* $\bar{\varepsilon}$, *a curve of class* C^1 *such that* $\xi(0) = p'$ *and* $\dot{\xi}(0) = v$. *Then for every neighborhood* N *of* p *in* S *the image* $F(N)$ *contains a neighborhood of the arc* $\{\gamma(t) : 0 < t < \varepsilon\}$ *in* S' *for some* $\varepsilon \in]0, \bar{\varepsilon}[$.

Hence, if one can prove the chain rule, one gets a good theory.
So far, I have not been able to prove the chain rule in full generality. All I know is the following weaker results:

I. The chain rule is true for **polyhedral cones**, but the proof that I have is very long and technical.

II. The following disjunction is true
   either
   (a) the chain rule does **not** hold for general closed convex cones,
   or
   (b) the chain rule for general closed convex cones is true but very hard to prove,

or

(c) I am very stupid.

The proof of the chain rule for polyhedral cones is very long and technical, and does not seem to generalize to arbitrary closed convex cones.

In any case, one can of course define the notion of "polyhedral local diffcone", and the theory works perfectly for local polyhedral diffcones.

CONCLUDING QUESTIONS:

A. Can one have a chain rule for PIGDS on arbitrary diffcones? (Answer: *I don't know.)*

B. Does it matter? (Answer: *for the Maximum Principle without constraints it certainly does not, because the only cones that are used are orthants, but for a maximum principle with state space constraints it probably will matter a lot. But this is just a guess.*)